

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2002-529981

(P2002-529981A)

(43) 公表日 平成14年9月10日(2002.9.10)

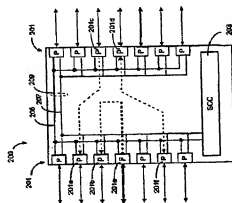
(51) Int.Cl. ⁷	識別記号	F I	テマート(参考)
H 0 4 L 12/56		H 0 4 L 12/56	F 5 K 0 2 6
H 0 4 M 3/00		H 0 4 M 3/00	A 5 K 0 3 0
H 0 4 Q 3/52		H 0 4 Q 3/52	A 5 K 0 6 1
			C 5 K 0 6 9
3/545		3/545	
審査請求 有 予備審査請求 有 (全 34 頁)			
<p>(21) 出願番号 特願2000-580396(P2000-580396)</p> <p>(86) (22) 出願日 平成11年8月17日(1999.8.17)</p> <p>(85) 翻訳文提出日 平成13年4月27日(2001.4.27)</p> <p>(86) 国際出願番号 P C T / U S 9 9 / 1 8 9 0 7</p> <p>(87) 国際公開番号 W O 0 0 / 2 7 1 3 6</p> <p>(87) 国際公開日 平成12年5月11日(2000.5.11)</p> <p>(31) 優先権主張番号 0 9 / 1 8 3 , 3 9 6</p> <p>(32) 優先日 平成10年10月29日(1998.10.29)</p> <p>(33) 優先権主張国 米国 (US)</p>			
<p>(71) 出願人 マックデータ コーポレーション アメリカ合衆国、コロラド、ブルームフ ールド、インターロッキン パークウェイ 310</p> <p>(72) 発明者 ネルソン、ジェフリー、ジェイ アメリカ合衆国 コロラド、ルイスビル、 サンランド ストリート 415</p> <p>(72) 発明者 ジェソップ、ケン、エス アメリカ合衆国 コロラド、アーバ、 キプリング ストリート 6432</p> <p>(74) 代理人 弁理士 浅村 皓 (外3名)</p>			

最終頁に続く

(54) 【発明の名称】 高性能クロスバー・スイッチ

(57) 【要約】

複数の入力/出力 (I/O) ポート (201) と、前記 I/O ポートの第1のものを前記 I/O ポートの第2のものへつなぐようにプログラムされて動作するクロスバー・デバイス (203) とを有するスイッチ (200)。複数のポート要求コントローラ (PRC) (300) は、各 PRC が前記 I/O ポートの1つに付随するようにつながれる。複数のシリアル要求バス (205) は、各シリアル要求バスが各 PRC をそれに付随するポートへつなぐように配置される。複数のシリアル応答バス (207) は、各シリアル応答バスが各 PRC をそれに関連する PRC へつなぐようにつながれる。動作時には、シリアル要求および応答バスは独立してノン・ブロッキング的に動作して接続およびクリア要求を並列的に処理する。



【特許請求の範囲】

【請求項1】 スイッチであって、複数の入力/出力 (I/O) ポート、前記 (I/O) ポートの内の第1の (I/O) ポートを第2の (I/O) ポートへつなぐようにプログラムされて動作するクロスバー・デバイス、複数のポート要求コントローラ (PRC) であって、各々が前記 (I/O) ポートの1つに関連している複数のポート要求コントローラ (PRC)、複数のシリアル要求バスであって、各々が各コントローラ (PRC) をそれに関連するポートへつないでいる複数のシリアル要求バス、および、複数のシリアル応答バスであって、各々が各コントローラ (PRC) をそれに関連するコントローラ (PRC) へつないでいる複数のシリアル応答バス、を含むスイッチ。【請求項2】 請求項1記載のスイッチであって、更に、ポート毎にビジー状態情報を保持するための記憶装置を含むビジー・ビット論理ユニットを含むスイッチ。【請求項3】 請求項1記載のスイッチであって、更に、関連する (I/O) ポートが最後に接続された (I/O) ポートの1つを示す識別子を保持するための記憶装置を含む、各 (I/O) ポートに関連するラスト・コネクテッド・ポート機構 (last connected port mechanism) を含むスイッチ。【請求項4】 請求項2記載のスイッチであって、前記ビジー・ビット論理が、ポート毎に接続要求を受信するための要求インタフェース、ポート毎に接続要求に対する応答を発行するための応答インタフェース、および、複数のコントローラ (PRC) の各々に対して、ポート毎に現在の利用可能状態を提供するための状態インタフェース、を含んでいるスイッチ。【請求項5】 請求項2記載のスイッチであって、前記ビジー・ビット論理ユニットが、更に、前記複数のコントローラ (PRC) の各々に対してそのコントローラ (PRC) に関連するポートに最後に接続された (I/O) ポートについての情報をポート毎に提供するためのラスト・コネクテッド・ポート・インタフェースを含むスイッチ。【請求項6】 請求項4記載のスイッチであって、各コントローラ (PRC) が、前記シリアル要求バスおよび前記シリアル応答バスへつながれたポート・インタフェース、前記ビジー・ビット論理の前記要求インタフェースに対して接続要求を送信するための要求インタフェース、各ビジー・ビット論理ユニットの前記応答インタフェースから接続要求に対する応答を受信するための応答インタフェース、および、前記ビジー・ビット論理ユニットからポート毎に現在の利用可能状態を受信するための状態インタフェース、を含んでいるスイッチ。【請求項7】 請求項1記載のスイッチであって、前記クロスバー・デバイスは、前記複数の (I/O) ポートの内の第1の (I/O) ポートを、該第1の (I/O) ポートによって指定される他の任意の (I/O) ポートへ接続するようにプロ

グラムされて動作するスイッチ。【請求項8】 請求項1記載のスイッチであって、前記クロスバー・デバイスは、前記複数の (I/O) ポートの内の第1の (I/O) ポートを他の (I/O) ポートの1組へつなぐようにプログラムされて動作するようになっており、ここで、前記他の (I/O) ポートの前記1組が前記第1の (I/O) ポートによって指定されるスイッチ。【請求項9】 請求項1記載のスイッチであって、前記シリアル要求バスが前記シリアル要求バスと独立して動作するスイッチ。【請求項10】 クロスバー・スイッチを制御するための方法であって、複数の入力/出力 (I/O) ポートを提供するステップ、前記複数の (I/O) ポートの内の第1の (I/O) ポートで第1の接続要求を発生するステップ、前記複数の (I/O) ポートの内の第2の (I/O) ポートで第2の接続要求を発生するステップ、および、前記第1および第2の接続要求を同時に処理するステップ、を含む方法。【請求項11】 請求項10記載の方法であって、前記同時に処理するステップが更に、前記第2の接続要求を発生する前記ステップのあとで、前記第1接続要求に対する応答を発生するステップ、を含んでいる方法。【請求項12】 請求項11記載の方法であって、更に、前記第1の接続要求に対する前記応答を受信する前に、前記複数の (I/O) ポートの前記第1の (I/O) ポートで第3の接続要求を発生するステップを含む方法。【請求項13】 請求項10記載の方法であって、前記複数の接続要求が互いに非同期である方法。【請求項14】 請求項11記載の方法であって、前記複数の接続要求および応答がノン・ブロッキング的に発生される方法。【請求項15】 請求項10記載の方法であって、更に、前記第1の接続要求に応答して、前記第1接続要求を満足する接続が既に前記クロスバー・スイッチ中にセットアップされているかどうかを示す信号を発生するステップ、および、前記第1の接続要求を満足するように既にセットアップされている前記接続を利用するステップ、を含む方法。【請求項16】 請求項10記載の方法であって、前記第1の接続要求が、前記接続要求に関する宛先ポートである複数の (I/O) ポートを指定しており、更に、前記第1の接続要求中に指定された前記宛先ポートの各々の利用可能性を同時に決定するステップ、を含む方法。【請求項17】 請求項16記載の方法であって、前記第1の接続要求が、前記第1のポートと前記複数の宛先ポートとの間にマルチキャスト接続を確立するための要求を含んでいる方法。【請求項18】 請求項16記載の方法であって、前記第1の接続要求が、前記第1のポートと前記複数の宛先ポートとの間にブロードキャスト接続を確立するための要求を含んでいる方法。【請求項19】 第1のポートを、該第1ポートによって発生される要求に答えて少なくとも1個の他のポートへ接続するためのクロス

バー・スイッチ用コントローラであって、前記要求を転送するために前記第1ポートへつながれたシリアル要求バス、前記第1ポートに関連し、前記要求を受信するための前記シリアル要求バスにつながれたポート要求コントローラ（PRC）、および前記第1ポートおよびその他の各ポートに関連し、前記コントローラ（PRC）へつながれて前記宛先の利用可能状態を表示する第1のビジー・ビット論理区分、を含むコントローラ。

【請求項20】 請求項19記載のコントローラであって、更に、前記少なくとも1個の他のポートに関連する第2のビジー・ビット論理区分であって、前記コントローラ（PRC）へつながれて前記関連ポートの利用可能性を表示する第2のビジー・ビット論理区分、を含むコントローラ。【請求項21】 請求項19記載のコントローラであって、更に、前記ソース・ポートへつながれ、前記コントローラ（PRC）へつながれてポートの利用可能状態を通信するためのシリアル応答バスを含むコントローラ。【請求項22】 請求項19記載のコントローラであって、更に、前記コントローラ（PRC）へつながれ、更に各々の他のポートへつながれて、前記ソース・ポートへ最後につながれた前記他のポートの任意のポートを表示するためのラスト・コネクテッド・ポート・バスを含むコントローラ。【請求項23】

通信システムであって、複数のネットワーク・デバイスであって、各デバイスが外部デバイスと通信するためのノードを有している複数のネットワーク・デバイス、複数の入力/出力（I/O）ポートを有するスイッチ、前記複数のノードの各々を前記スイッチの前記（I/O）ポートの1つへつなぐ通信チャネル、前記スイッチ内において、選ばれたソースI/Oポートを選ばれた宛先I/Oポートへつなぐようにプログラムされて動作するクロスバー・デバイス、複数のポート要求コントローラ（PRC）であって、各コントローラ（PRC）が前記複数の（I/O）ポートの選ばれた1つに関連している複数のポート要求コントローラ（PRC）、複数のシリアル要求バスであって、各シリアル要求バスが選ばれた1つのコントローラ（PRC）をそれに関連するポートへつなぐ複数のシリアル要求バス、および複数のシリアル応答バスであって、各シリアル応答バスが選ばれた1つのコントローラ（PRC）をそれに関連するPRCへつなぐ複数のシリアル応答バス、を含む通信システム。【請求項24】 請求項23記載の通信システムであって、前記宛先I/Oポートが、前記シリアル要求バスの1つの上を送信される接続要求に基づいて選ばれた通信システム。

【発明の詳細な説明】

【0001】（発明の分野） 本発明は一般にクロスポート・スイッチ用のコントローラに関するものであって、更に詳細には、ファイバ・チャネル・スイッチ中で高速に接続をセットアップするための方法および構

造に関する。【0002】（関連背景） ファイバ・チャネルは、サーバ、記憶システム、ワークステーション、スイッチ、およびハブの間の双方向的なポイント・ツー・ポイント通信用に設計された高性能なシリアル相互接続の標準である。これは他のリンク・レベルのプロトコルと比べて多様な特長を有し、それらの中には効率および高性能、スケラビリティ、簡便さ、使用および導入の容易さ、汎用的な高レベル・プロトコルのサポートが含まれる。【0003】 ファイバ・チャネ

ルは、ポート間に接続を確立するために“ファブリック（fabric）”として知られるトポロジーを採用している。ファブリックは、スイッチの構成に制限を設けずに複数のデバイスを相互接続するスイッチ網である。ファブリックは、ポイント・ツー・ポイントおよびアービトラート（調停）・ループの両トポロジーが混在したものを含むことができる。【0004】 ファブリック・チャネルでは2つのノード間にチャネルが確立され、チャネルの主要な役割は1つのポートから別のポイントへ短い待ち時間で高速にデータを転送することであり、ハードウェア的には簡単なエラー検出のみを実行する。ファイバ・チャネル・スイッチは、複数の同時的なポイント・ツー・ポイント接続を確立することによって、柔軟な回路/パケット交換トポロジーを提供する。これらの接続は、接続された末端のデバイスや“ノード”ではなく、スイッチあるいは“ファブリック要素”によって管理されるため、ファブリック通信管理はデバイスの観点からみて大幅に簡素化される。【0005】 ファイバ・チャネルへ接続するために、ファブリック・デバイスにはファブリック接続を管理するためのノード・ポートあるいは“N_ポート”が含まれる。N_ポートは、ファブリック・ポートあるいは“F_ポート”を有するファブリック要素（例えばスイッチ）への接続を確立する。ファブリックへとりつけられるデバイスは、N_ポートとF_ポートとの間の接続を管理するだけの知能（インテリジェンス）を要求される。ファブリック要素は、ルーティング、エラー検出および回復、および同様な管理機能処理する知能を含む。【0006】 スイッチはマルチポート・デバイスであって、各ポートはそれ自身とそれにつながれたシステムとの間の簡単なポイント・ツー・ポイント接続を管理する。各ポートは、サーバ、周辺機器、I/Oサブシステム、ブリッジ、ハブ、ルータ、あるいは別のスイッチへつなぐことができる。スイッチは1つのポートから接続要求（リクエスト）を受信し、別のポートへの接続を自動的に確立する。マルチポート・ファイバ・チャネル・スイッチを通して、複数の呼び出しやデータ転送が同時に発生する。交換技術のキーとなる特徴は、その“ノン・ブロッキング”性であり、一旦スイッチを通る接続が確立されれば、その接続によって提供される帯域は共有されない。【0007】 ファイバ・チャネル・ファブリ

ックでのデータ接続は短時間になる傾向があり、そのため、スイッチを通る接続のセットアップに時間の掛かるスイッチはそのスイッチの有効な帯域を大きく制限することになる。接続要求が行なわれてその接続がセットアップされるまでに複数のマシンのクロック・サイクルを要するのが一般的である。1つのポートでデータ接続の要求が行なわれてからその接続が実際にセットアップされるまでの間は、それらの2つの該当ポート間でデータは転送できない。【0008】従来の方式では、

マルチポート・スイッチは各種ポートからの接続要求を逐次的に処理したため、もし1つの接続要求が発行されれば、その最初の接続要求が処理されるまでは後続の接続要求は抑制された。この方式ではポートとクロスバー制御論理との間で接続要求を送信するために複数のポートによって共有される共通制御バスが使用されていた。このため、1つの障害によって複数のポートが影響を蒙った。更に、接続セットアップは各ポートに関する接続要求を処理する単一のパイプラインによって処理された。スイッチ当たり68ポート（あるいはそれ以上）までは、接続セットアップの間にバス接続によって引き起こされる待ち時間はスイッチ全体のスループットにとって重大なボトルネックとなってきた。【0009】従来のシステムの別の制限は、接続要求を発行するポートがそれが接続しようとしている宛先ポートがビジーであるかどうかについて何も知らないということである。このため、要求を発行しているポートは複数のクロック・サイクルを要する接続要求処理をすべて行わなければならないが、そのあとで、その接続が失敗に終わるということもあった。しかし、目的ポートのビジー状態は、接続要求処理が開始される前に設定されるはずである。ポートのビジー状態をそのビジーなポートへの接続を要求しているポートへ知らせることによって、要求を発行しているポートがその要求をあきらめるか、遅らせることでスイッチ資源の浪費を減らすことができるようにする必要がある。【0010】（発明の概要）要約すれば、本発明は、複数の入力/出力（I/O）ポートと、前記I/Oポートの第1のもので前記I/Oポートの第2のものでつながるようにプログラムできるクロスバー・デバイスとを有するスイッチを含む。複数のポート要求コントローラ（PRC）が、各PRCがI/Oポートの1つに関連若しくは付随するようになされる。複数のシリアル要求バスが、各シリアル要求バスが各PRCをそれに関連若しくは付随するポートへつなぐように配置される。複数のシリアル応答バスが、各シリアル応答バスが各PRCをそれに関連するPRCへつなぐようにつなされる。【0011】別の態様において、本発明は複数の入力/出力（I/O）ポートへつながれたクロスバー・スイッチを制御するための方法を含む。前記複数のI/Oポートの第1のもので第1の接続要求が発行され、前記複数のI/Oポートの第2のもので第2の

接続要求が発行される。これらの第1および第2の接続要求は同時に処理される。前記シリアル要求およびシリアル応答バスは独立してノン・ブロッキング的に動作して、接続およびクリア要求を並列的に処理する。【0012】（好適な実施の形態の詳細な説明）図1は本発明に従う方法およびシステムを実現するための一般化されたファイバ・チャネル環境を示す。図1は複数のデバイスと、デバイス間にラインをつなぐことによって形成された接続とを示している。各デバイスは1または複数のポートを含む。ファイバ・チャネル中で、これらのポートは一般にノード・ポート（N__ポート）、ファブリック・ポート（F__ポート）、および拡張ポート（E__ポート）に分類される。ノード・ポートはサーバ101、ディスク・アレイ102および103、およびテープ記憶デバイス104などのノード・デバイス中に位置する。ファブリック・ポートはスイッチ106およびスイッチ107などのファブリック・ポート・デバイス中に位置する。オプションとして、アービトレーテッド・ループ・ネットワーク108のようなローカル・エリア・ネットワークはファブリックなアービトレーテッド・ループ・ポート（FL__ポート）を使用してファブリック・リンクである。【0013】図1に示すデバイスはラインで示す“リンク”によって接続される。2つのノード・ポート間に1または複数のリンクを用いてチャネルが確立できる。例えば、例え図1に示すようなサーバ101とディスク・アレイ102との間のチャネルのように、単一のリンクを用いて直接的あるいはポイント・ツー・ポイントのチャネルが確立される。この場合、各N__ポートは、他のデバイスのN__ポートとの間のポイント・ツー・ポイント接続を管理する。多重リンクを用いて回路交換されるチャネルもまた、スイッチ106または107を用いても提供できる。例えば、サーバ101中のN__ポートはスイッチ106を通してディスク・アレイ103のN__ポートとの間にチャネルを確立できる。この場合、各N__ポートはスイッチ106のF__ポートへ接続する。スイッチ106はまた拡張ポートまたはE__ポートを含み、それを用いてスイッチ107の別のE__ポートへチャネルを確立できる。インター・スイッチ・リンク（ISL）（すなわち、スイッチを含むリンク）を用いば、各N__ポートはファブリック中の他の各N__ポートへ1または複数の経路を通ってつなぐことができる。スイッチ107はまた、アービトレーテッド・ループ108へのリンクをサポートするファブリック・ループ・ポート（FL__ポート）を含む。ループ108で、ハブ109は、スイッチ107へのリンクをサポートするFL__ポートとともに、ループ・サーバ110およびワークステーション112および113へのリンクをサポートするノード・ループ・ポート（NL__ポート）を含んでいる。デバイス111、112、および113は従来の直接接続またはネットワーク接続を通し

てファブリックへつながるユーザ端末、パーソナル・コンピュータ、あるいはワークステーションを表している。

【0014】 図2は、本発明に従う高性能なクロスバー交換の方法および装置を採用する特別なスイッチ200を示している。各ポート201（図2ではPと表記している）は外部デバイスとの間で双方向のデータ通信をサポートする入力/出力（I/O）ポートである。各ポートには任意の利用可能な設計および技術による送信および受信回路（図示されていない）が含まれる。しかし、本発明の目的のためには、スイッチ200は複数の専用入力ポートおよび複数の専用出力ポートで以って等価的なものが構成できる。【0015】 スイッチ200は、破線で図示するように構成可能なクロスバー・デバイスを含み、これは任意の1つのポートを別の任意のポートへつなぐようにプログラムできる。クロスバー・デバイスは完全なクロスバーであることが好ましいが、部分的なクロスバー・デバイスであっても、本発明の教えるところから離れることなく等価なものとして置換でき、その場合の性能に対する影響は予測可能である。接続は、図2の矢印の向きによって示されるように、目的に応じて一方向的（すなわち、半二重）であっても双方向的（すなわち、全二重）であってもよい。外部データを受信するポートは「ソース・ポート」と呼ばれ、他方、外部の宛先ヘッダを送信するポートは「宛先ポート」と呼ばれる。従って、スイッチ202中でソース・ポートから宛先ポートへ接続が形成される。図2中で接続例はポート間に破線で示されており、矢印の向きは宛先ポートを向いている。【0016】 シリアル・クロスバー・コントローラ（SCC）203は、要求部205および応答部207を含む複数の独立した接続バス209によって各ポートへつなぐがれて動作する。要求部205は複数の独立したシリアル要求バスを含み、その各々は1つのポート201を1つのPRC305へつなぐ接続ラインを含んでいる。同様に、応答部207は複数の独立したシリアル応答バスを含み、その各々は1つのPRC305を1つのポート201へつなぐ接続ラインを含んでいる。以下に詳細に説明するように、各シリアル要求バスと各シリアル応答バスとは独立しているため、要求と応答を同時に処理することができ、それによって待ち時間を減らすことができる。ここで用いる「同時に」という表現は「1つのクロック・サイクル中で処理される」ことを意味する。言い換えれば、同時要求はノン・ブロッキング的に処理され、そのため未決要求は未決要求を発行しているポート201あるいはスイッチ200中の他のポート201のいずれかからの他の要求の発行および処理を妨害しない。確立することのできなかつた要求は並列的に処理されるため、従来の逐次的な接続バス実行が要したような接続バスの帯域の消費はない。【0017】 SCC203は、各ポート201から接続要求を受信し、クロスバー

・デバイス中で構成可能な接続をプログラムする。動作時には、ソース・ポートは外部デバイスからデータ・パケットを受信する。各ポート中の論理回路は、受信したデータが接続セットアップを要求すること、およびSCC203への要求バス205上へ接続要求を行うことを検出する。SCC203は、接続を行うことができるあるいは接続ができないことを表示する接続状態情報によって応答バス207上で応答する。本発明に従えば任意の時点で2以上のポートが接続要求を受信できるため、SCC203は好ましくは、待ち時間を短くするように接続要求を同時に（すなわち、並列的に）処理する。【0018】 図3はSCC203の主要部品をブロック図で示す。接続バス209（図2に示されている）の要求部205は複数の接続ラインを含んでおり、各接続ラインはシリアル要求バス205であり、それはポートの1つ201（図2に示されている）から1つのポート要求コントローラ（PRC）305との間のポイント・ツー・ポイント接続を行う。SCC203は好適な実施の形態において、各ポート201に対して1つのPRC305（N個のI/Oポートを有するスイッチに対してPRC₀ないしPRC_Nと各づけられている）を含む。実際には、SCC203は、従来のポート管理機能を実行するために、付加的な論理、メモリ、および信号処理デバイスを任意の数だけ含む。これらの付加的な論理デバイスは、本発明の特徴および動作の理解を助けるために図示されていない。【0019】 各PRC305は接続バス209（図2に示されている）の要求部205内の接続ラインの1本（同時要求の場合には2本以上）上でシリアル・ビット・ストリームを受信する。各PRQ305は単一のポート201に関するポート接続を管理（マネージング）することに専念しており、個別ポート201間に接続セットアップを実行するためにシリアル・ビット・ストリームを処理するように構成される。特別な応用に従っていくつかのタイプの要求が定義できる。実施例では、要求は、次のような4つの異なるタイプの要求：「クリアのみ」、「全二重接続」、「クリア付きの半二重接続」、そして「クリアなしの半二重接続」を表す符号化情報を含む。各PRC305は要求バス205の接続ラインをスキャンしてアクティブな要求を検出する。特別なPRC305によって要求バス205上で要求が検出された場合には、PRC305はその要求タイプを復号して要求のチェック・ビットを検証する。【0020】 送信区分303のビジー状態（ステータス）は各個別ポート201についてビジー・ビット論理307によって保持される。ビジー・ビット論理307は本質的に複数の状態機械であって、各状態機械は付随するポート送信機クリアまたはビジー状態を保持する。典型的な実施例では、ビジー・ビット論理307はスイッチ200中の各ポート201に対してエンタリーを有する集中化されたメモリ・アレイであ

る。各ポート201に対して、付随するビジー・ビット論理は要求バス205上での他のポート201からの要求を受け入れ、応答バス207上へ応答を送信して、要求を発しているPRC305へその状態を表示する。【0021】ポートがビジーであるのはその送信機が使用中である場合であり、そうでなければ利用できる。ビジー・ビット論理307はPRC305からの要求メッセージを受信し、適切なPRC305へ応答メッセージを送信するためのインタフェース論理を含む。ビジー・ビット論理307のインタフェース論理はポイント・ツー・ポイントであり、同じポート201または異なるポート201のいずれかへ多重要求の同時発生を許可する。多重接続が異なるポートに対して行われる場合には、それぞれ対応するPRQ305はクロスバ一接続論理(CCL)302に対して調停を行って接続を確立する。【0022】CCLでの接続セッアップには多くのクロック・サイクルが消費され、同じ接続資源を競うポート間の調停(アービトレーション)は更に待ち時間(レイテンシー)を加えるであろう。好適な実施形態では、ビジー・ビット論理307は、要求を発しているポートが最後にその宛先につながっていたかどうかを判定するためのデマイスを含む。各PRC201は、それがその要求中で指定されたその宛先ポートへの接続を要求する最後のポートであったかどうか(すなわち、要求された宛先ポートへ接続がセッアップされたままであるかどうか)を知っている。そのような場合には、既に確立されているクロスバ一接続を用いることができ、新しいクロスバ一接続を行う必要はなく、効率は増大する。【0023】図4は、接続バス209上を転送されるシリアル接続メッセージ例を示している。図4に示される接続および応答バス間には支配的若しくは強制的なタイミング関係は必要でない。言い換えれば、図4は2つの同時要求メッセージを1つの同時応答メッセージと一緒に示しているが、これらのメッセージは動作的には互いに同期していないし、同時に発生していてもそうでなくとも構わない。要求バス205および応答バス207中の各接続ラインは一時に1つのメッセージだけをサポートできるが、任意の時点で任意数の接続ラインがメッセージを運ぶことはできる。【0024】各PRC305および各ポート201へシステム・クロック信号401は5.3MHzの信号である。図4は水平ライン上に205aおよび205bと名づけた2つの異なるタイプの要求を示しており、図4に示す要求が要求バス205上の接続ラインの1つでアサートされることを示している。同様に、図4は207と名づけた水平ライン上に1つの応答メッセージを示している。特別な例では、すべての応答が同じメッセージ形式を取るため、1つの応答メッセージだけを示して説明すればよい。しかし、本発明は、2つ以上のタイプのメ

ッセージを含む多様な応答メッセージを受け入れるように容易に適應できる。【0025】要求バス205上で、(図4でTYPE_0およびTYPE_1と名づけた)次の2クロック・サイクル間の接続ラインの状態(すなわち、アクティブか非アクティブか)は実際の要求タイプを識別する符号化を示している。この構成によって4つの要求タイプが許容され、より多くのクロック・サイクルを使用することによって非常に多数の符号化が提供でき、より多様な要求タイプが表示できるようになる。表1は要求タイプの符号化例と要求メッセージ中に割当てられるビット値とを示す。表1に示す符号化の割当てはほんの一例に過ぎず、本発明に従えば、より多いあるいは少ないビットを使用する他の符号化割当ても同様に可能であることを理解されたい。【0026】

【表1】

要求ビット	TYPE_0	TYPE_1	符号化要求タイプ
0	0	0	クリアのみ
0	1	0	クリアおよび非多重接続
1	0	0	非多重接続のみ
1	1	1	全多重接続

表1

【0027】図4で、ライン205a上の要求はクリアのみの型を要求を含んでおり、それは受信を行うPRC305に指示してビジー・ビット論理307の関連部分のビジー・ビットをクリアさせる(すなわち、クリア要求を発行するポート201の利用可能性を表示させる)。ライン205a上のクリアのみの要求に対して、接続ラインは図4のTYPE_0およびTYPE_1と名づけた同期間において非アクティブ状態にセッされる。特別な実施例では、ライン205a上のクリアのみの型は4クロック・サイクルだけが必要とする。第1のクロック・サイクルでは、ライン205a上の論理信号はアクティブ状態にセッされて、要求がアサートされていることを表示する。関連するPRC305はそれがつながれている各接続ラインをクロック・サイクル当たり少なくとも一度はスキャンして、アクティブ状態にある接続ラインを検出する。【0028】クリア要求は受信しているPRC305に指示してそれ自身のポートに関連するビジー・ビットをクリアさせるため、付加的なデータは必要ない。図示のように、クリア・メッセージには要求ビットがアクティブにセッされたあとで、3番目のクロック・サイクル中に追加されたCHECKビットが含まれる。CHECKビットはエラー検出のために使用される。特別な例では1つのCHECKビットだけが使用されているが、特別な応用で必要とされる場合には、利用可能なエラー検出/修正アルゴリズムを用いてエラー検出/修正を行えるように、2つ以上のCHECKビットを使用してもよい。このように、クリア要求メッセージは4クロック・サイクルの間に通信される。TYPE_0、TYPE_1、CONNECTおよび

びCHECKビットは要求メッセージ中のどの場所に配置しても等価である。【0029】ライン205b上に示す接続要求はクリアおよび半二重接続、半二重接続のみ、および全二重接続の各要求タイプのタイミングを示している。TYPE_0、TYPE_1、およびCHECKビットはライン205a上に示すクリアのみメッセージにおけるものと同じ機能をする。しかし、“接続メッセージ”とも呼ばれる接続を要求する要求メッセージ中で、TYPE_0およびTYPE_1ビットは表1に示すような所望の要求タイプを表示するために適切にセットされる。接続メッセージは複数のビット(例えば、ビットD0-D7)を含み、それらは接続を行うべき宛先ポートを指定する。利用可能なポート数を受け入れるために任意数のデータ・ビットを含めることができる。例えば、図4に示すような8ビットでは接続メッセージ中で識別できるポート数は2⁸あるいは256個となる。このように、1つのポートと他の256個のポートの1つとの間に接続を要求するシリアル接続メッセージは12クロック・サイクル間に送信できる。【030】図4のライン207はシリアル応答メッセージ例を示している。この特別な例では、各応答に5ビットが含まれ、それらは応答ラインがアクティブであることを知らせる第1ビット、3つのデータ・ビット、そして1つのCHECKビットである。PRC305は、応答フェーズの1クロック・サイクルの間に、ライン207をアクティブにセットすることによって応答を表示する。応答フェーズに続く3クロック・サイクルは接続要求の状態を示す。3つの応答ビットによって2³個のユニークな応答状態を表示できる。表2は応答状態の符号化例と応答メッセージ中に割当てられるビット値を示す。表2に示す符号化割当てではほんの一例に過ぎず、本発明に従えば、より多いあるいは少ないビットを使用する他の符号化割当てでも同様に可能であることを理解されたい。好ましい応答メッセージにはCHECKビットが含まれ、それは先に説明した要求メッセージ中のCHECKビットと同様に機能する。【0031】表2]

応答ビット	定義
000	宛先ポート半二重として接続
001	ソース・ポート半二重として接続
010	ソースおよび宛先ポート半二重として接続
011	ソース・ポート全二重として接続
100	宛先ポート全二重として接続
101	要求された宛先ポート半二重で、宛先ポート半二重として接続
110	要求された宛先ポート全二重
111	ソース・ポート全二重

【0032】SCC203をより完全に理解することとは、図5に示す特定の68ポート実施例を参照することによって可能である。図5は、ビジー・ビット・ユニット307内に68ビジー・ビット・セグメント501とともに68個のPRC305を含む68ポート・デバイスほんの一部を示している。各PRC305には1

つのポート201が付随する(図2に示されている)。同様に、各ビジー・ビット・セグメント501には1つのポート201が付随する。【0033】各PRC305には、特別な応用での必要に応じて、1または複数のビット幅を有する、クロスバー接続論理302(図3に示されている)への接続が含まれる。好ましくは、各PRC305は任意の利用可能なクロスバー・アドレス指定法を利用してCCL302へ多重ライン並列接続を行う。【0034】各PRC305は1本のビジー・ビット・クリア(BB_CLR)ラインを含んでおり、それはそのポートに付随するビジー・ビット・セグメント501へつながれる。ポートの利用可能性はBB_CLRライン上へアクティブ信号をアサートすることによってビジー・ビットをクリアすることによって表示される。PRC305によって検出された要求がそれ自身のビジー・ビットをクリアする要求である場合は、PRC305は要求メッセージの検証ビットをチェックしたあとで、ビジー・ビット・クリア(BB_CLR)信号をアサートする。一旦ビジー・ビットがクリアされれば、ポート201は他の接続のための宛先ポートとして利用できるようになる。【0035】任意の接続要求に対して、PRC305は接続要求のシリアル・ビット・ストリームから宛先ポート番号を復号し、チェック・ビットを検証する。PRC305がその要求が有効であることを判断した場合には、PRC305はBB_REQラインの1つにアクティブ信号をアサートすることによって、指定された宛先ポートを制御しようと試みる。BB_REQバスは図5の68ポートの例では68本の接続ライン(BB_REQ[67:0])を含む。各BB_REQラインは1つのPRC305(例えば、図5のPort_0要求コントローラ305)を各ビジー・ビット・セグメント501とつなぐ(例えば、BB_0ないしBB_67)。特別な例では、PRC305は、図4に示すように、1つのクロック周期の間に復号された宛先ポートのBB_REQ信号をアサートする。【0036】各ビジー・ビット・セグメント501は各PRC305からのBB_REQ信号を受信するためのBB_REQインタフェースを含む。各ビジー・ビット・セグメント501はまた、各PRC305へ散開する1本のビジー・ビット・ライン(例えば、BB[0])とともに、各PRC305へつながれる応答ライン(例えば、BB_RSP[0]-BB_RSP[67])を含む。このように、各PRC305は各ビジー・ビット・セグメント501のビジー/利用可能状態を継続的に監視しており、従って各ポート201の利用可能性を知っている。要求を発するPRC305は、アドレス指定されるビジー・ビット・セグメント501が応答するのを待つ必要がない。しかし、BB_REQライン上へ信号をアサートを送信する前にPRC305がビジー/利用可能状態を知っているとしても、BB_REQ信号はチェ

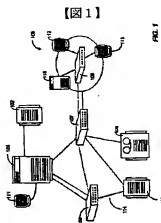
ックされずにアサートされて、宛先ポートのビジー/利用可能状態を評価する間のBB_REQ信号の妨害に付随する待ち時間を回避するようにしている。【0037】

接続要求メッセージの最後に、接続要求を受信しているPRC305は適切なBBライン（すなわち、図5のBB[67:0]ラインの1本）およびそのBB_RSPライン（すなわち、図5のBB_RSP[67:0]ラインの1本）をスキャンする。要求される宛先ポートに関してBBがセットされているか、あるいはBB_RSPがセットされていなければ、接続を行なうことはできず、PRC305は要求を発しているポート201に対して宛先ポートビジー・メッセージを戻す。接続が確立された（すなわち、BBがクリアで、BB_RSPがセット）場合には、要求を発しているPRC305はCCL302に対してスイッチ要求を発し、要求を発しているポート201に対して適切な応答メッセージを戻す。【0038】オプションの好適な実施例では、ビジー・ビット論理セグメント501に「ラスト・コネクテッド・ポート」情報を含めることによってCCL302へのスイッチ要求発生を最小化している。各ビジー・ビット論理セグメント501は図5の「0-67」と名づけられたラインを含み、それは関連するPRC201に対して前の接続要求の結果としてCCL302によって接続が既にセットアップされていることを表示する。要求されるポートへの接続が許可された場合には、「ライン上の表示」を利用して、クロスバーが既に構成されているかどうか、あるいは再構成が要求されているかどうか決定される。再構成が要求されない場合には、要求を発しているPRC201はCCL302へのスイッチ要求発生工程をパスすることができる。【0039】付加的な特徴として、本発明はブロードキャストおよびマルチキャスト・グループをサポートするように容易に適応できる。ブロードキャストおよびマルチキャストでは、1個のソース・ポート201が複数の宛先ポート201へ接続を要求し、各宛先ポート201が先にデータを受信するようになる。一時に1つだけの接続セットアップを行う従来のスイッチ構成に対して、マルチキャストおよびブロードキャスト構成に対して望ましくない待ち時間が加えられる。【0040】実施例では、宛先ポートのアドレス空間の一部（例えば、図4の8ビットD0-D7で定義されるアドレス空間）がマルチキャスト/ブロードキャスト・グループを表示するために確保される。あるいは、マルチキャスト/ブロードキャスト要求を表1に示すように特殊な要求型として指定して、ユニークな符号化を判定してもよい。特別な実施例では、16進数「0×F0」でない16進数「0×FF」の範囲にある宛先ポートアドレスを含む接続要求メッセージによってマルチキャストまたはブロードキャスト接続を指定する。【0041】図6は、マルチキャスト構成を処理する集中マルチキャスト・ユニット

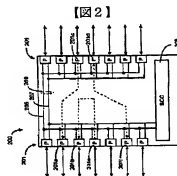
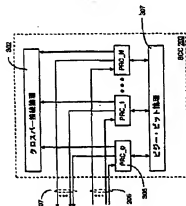
（図示されていない）または各PRC305中に組み込むことのできる構成機例を示す。システム管理者は（手動または自動）スイッチ200を16個のマルチキャスト/ブロードキャスト・グループまたはセットに予め構成してある。各マルチキャスト/ブロードキャスト・グループには構成レジスタ600中のエントリー601が付随する。各グループまたは組は、16進数「0×F0」でない16進数「0×FF」の範囲にある16個のアドレスの1つに対応する。各マルチキャスト/ブロードキャスト・グループは、要求を発している1本のソース・ポートへ同時接続できる1組の宛先ポートを含む。マルチキャスト/ブロードキャスト・グループへのポート割当てはマルチキャスト/ブロードキャスト・グループ内の特別なポートの利用者およびデバイスの特殊な要求に合致するように行なわれるのが普通である。【0042】構成レジスタ600中の各エントリー601は好適な実施例においてイネーブル・ビットを有しており、それによってシステム管理者は各マルチキャスト/ブロードキャスト・グループを独立的に選択的に許可することができる。例えば、エントリー601に記憶された論理「高レベル」はそのマルチキャスト/ブロードキャストを許可し、他方、論理「低レベル」はそのマルチキャスト/ブロードキャスト・グループを禁止する。特定のエントリー601を禁止することによって、対応するアドレス（すなわち、「0×F0」ないし「0×FF」の範囲にある1つのアドレスを従来のユニキャスト・アドレスとして使用できる。【0043】グループ・メンバーを指定するためにビット・マスク602を使用することも容易である。ビット・マスク602には各エントリー601、従って各マルチキャスト/ブロードキャスト・グループが付随している。ビット・マスク602にはマルチキャスト/ブロードキャスト・グループに含まれるスイッチ200中の各ポート201に対するエントリーが含まれる（例えば、68ポート・スイッチに対して68までのエントリー、あるいは256ポート・スイッチに対して256までのエントリー）。各エントリーはスイッチ・ポートの1つに対応し、1ビットの識別子を含み、それによって対応するスイッチ・ポートが特別なマルチキャスト/ブロードキャスト・グループに含まれるかどうかを表示する。【0044】要求を発しているPRC201はビット・マスク602によって指定された宛先ポートの各々に対して接続要求を発生する。多重並列接続要求は先にユニキャスト接続に関して説明したのと同じ機構およびプロトコルを使用する。本発明はポート毎に、独立した要求/応答資源を含むため、多重要求が同時に処理されて、その結果、BB、BB_RSP、および「上」に要求されたポートの利用可能状態を表示する複数の応答が得られる。【0045】多くの場合、要求される宛先のすべてが利用できるわけではない。このことは、マルチキャスト/ブロードキ

キャスト・チャンネルが、接続が確立できない場合（例えば、クラス2またはクラス3のサービス）に、パケットが切り捨てられたり、廃棄されたりするタイプのファイバ・チャンネル・マルチキャスト/ブロードキャスト仕様である場合には許容できる。要求を発しているPRC305はCCL302と一緒にスイッチ要求を開始し、利用可能な宛先ポート201への接続をセットアップする。そして要求を発しているPRC305に関連するポート201に対してメッセージを展して、マルチキャスト/ブロードキャスト要求の中でどの宛先ポートで調停に成功したかを表示する。【0046】本発明は複数の同時接続要求の処理を可能とするため、マルチキャストおよびブロードキャスト接続セットアップに関する待ち時間は大幅に短縮される。任意のマルチキャストまたはブロードキャスト・グループに対して、要求はすべて同じアービトラージ・ウィンドウの中で実行され、利用可能なメンバー・ポートはクロスバー接続論理

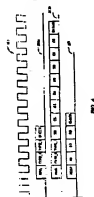
203に対してラッチされる。【0047】本発明について特定の実施の形態に関して説明してきたが、本開示はほんの一例に過ぎず、請求の範囲に定義される本発明の精神およびスコープから離れることなく部品の組合せおよび配置についての数多くの変更が当業者には明らかであろうことを理解されたい。【図面の簡単な説明】【図1】本発明に従う方法およびシステムを実現する一般化されたファイバ・チャンネル環境図。【図2】本発明に従う方法および装置を実現するスイッチのブロック図。【図3】本発明に従うシリアル・クロスバー・コントローラの詳細図。【図4】本発明に従うシリアル要求およびシリアル応答メッセージのタイミング図。【図5】本発明に従うポートおよびビジー・ビット論理の特別実施例。【図6】本発明の実施の形態に従ってマルチキャストおよびブロードキャスト要求を処理する構成機構例。



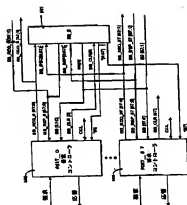
【図3】



【図4】



【図5】



【図6】

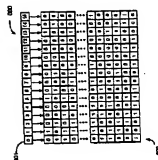


FIG. 6

【手続補正書】特許協力条約第34条補正の翻訳文提出書

【提出日】平成12年11月30日(2000. 11. 30) 【手続補正1】【補正対象書類名】明細書【補正対象項目名】請求項1【補正方法】変更【補正内容】

【請求項1】スイッチであって、複数の入力/出力(I/O)ポート、前記(I/O)ポートの内の第1の(I/O)ポートを第2の(I/O)ポートへつなぐようにプログラムされて動作するクロスバー・デバイス、

複数のポート要求コントローラ(PRC)であって、各々が前記(I/O)ポートの1つに関連している複数のポート要求コントローラ(PRC)、複数のシリアル要求コントローラ(PRC)をそれぞれに関連するポートへつなぐ複数のシリアル要求バス、および、複数のシリアル応答バスであって、各々が各コントローラ(PRC)をそれぞれに関連するポートへつなぐ複数のシリアル要求バス、を含むスイッチ。【手続補正2】【補正対象書類名】明細書【補正対象項目名】請求項3【補正方法】変更【補正内容】1【請求項3】スイッチであって、複数の入力/出力(I/O)ポート、前記(I/O)ポートの内の第1の(I/O)ポートを第2の(I/O)ポートへつなぐようにプログラムされて動作するクロスバー・デバイス、 複数のポート要求コントローラ(PRC)であって、各々が前記(I/O)ポートの1つに関連している複数のポート要求コントローラ(PRC)、複数のシリアル要求バスであって、各々が各コントローラ(PRC)をそれぞれに関連するポートへつなぐ複数のシリアル要求バス、 複数のシリアル応答バスであって、各々が各コントローラ(PRC)をそれぞれに関連するポートへつなぐ複数のシリアル要求バス、および、各(I/O)ポートに関連するラスト・コネクテッド・ポート機構であって、関連する(I/O)ポートが最後に接続された1つの(I/O)ポートを示す識別子を保持するための記憶装置を含むラスト・コネクテッド・ポート機構、を含むスイッチ。【手続補正3】【補正対象書類名】明細書【補正対象項目名】請求項5【補正方法

】変更【補正内容】【請求項5】スイッチであって、 複数の入力/出力(I/O)ポート、前記(I/O)ポートの内の第1の(I/O)ポートを第2の(I/O)ポートへつなぐようにプログラムされて動作するクロスバー・デバイス、 複数のポート要求コントローラ(PRC)であって、各々が前記(I/O)ポートの1つに関連している複数のポート要求コントローラ(PRC)、複数のシリアル要求バスであって、各々が各コントローラ(PRC)をそれぞれに関連するポートへつなぐ複数のシリアル要求バス、 複数のシリアル応答バスであって、各々が各コントローラ(PRC)をそれぞれに関連するポートへつなぐ複数のシリアル応答バス、および、ビジー・ビット論理ユニットであって、ポート毎にビジー状態情報を保持するための記憶装置を含み、更に、前記複数のコントローラ(PRC)の各々に対してそのコントローラ(PRC)に関連するポートに最後に接続された(I/O)ポートについての情報をポート毎に提供するためのラスト・コネクテッド・ポート・インタフェースを含む前記ビジー・ビット論理ユニット、を含むスイッチ。【手続補正4】【補正対象書類名】明細書【補正対象項目名】請求項22【補正方法】変更【補正内容】【請求項22】第1のポートを、該第1ポートによって発生される要求に応答して少なくとも1個の他のポートへ接続するためのクロスバー・スイッチ用コントローラであって、前記要求を転送するために前記第1ポートへつなぐれたシリアル要求バス、前記第1ポートに関連し、前記要求を受信するための前記シリアル要求バスにつなぐれたポート要求コントローラ(PRC)、前記第1ポートおよびその他の各ポートに関連し、前記コントローラ(PRC)へつなぐられて前記宛先の利用可能状態を表示する第1のビジー・ビット論理区分、および、前記コントローラ(PRC)へつなぐられ、更に各々の他のポートへつなぐられて、前記ソース・ポートへ最後につなぐれた前記他のポートの任意のポートを表示するためのラスト・コネクテッド・ポート・バス、を含むコントローラ。

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/JP99/18907
A. CLASSIFICATION OF SUBJECT MATTER IPC(8): H04Q 100, 300, 900; G06F 104, 1300, 1306A US CL.: 340/823.79, 825.8, 825.83; 395/275, 325, 425, 550; 710/1.33 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S.: 340/823.79, 823.8, 825.83; 395/275, 325, 425, 550; 710/1.33 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) WEST		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,555,543 A(GROHOSKI et al) 10 September 1996, col. 2, lines 15-19.	1-24
Y	US 4,605,928 A(GEORGIOU) 12 August 1986, col. 4, lines 25-30; fig. 3.	1-24
Y	US 5,392,422 A (HOEL et al) 21 February 1995, col. 7, lines 60-65.	1,4,6,9,10,19,23
Y	US 5,280,591 A (GARCIA et al) 18 January 1994, col. 5, lines 19-30; FIG. 4.	1,4,6,9,10,19,23
Y	US 5,754,120 A(ARGENTATI) 19 May 1998, col. 13, lines 30-32.	2
<input type="checkbox"/> Further documents are cited in the continuation of Box C. <input type="checkbox"/> See patent family member.		
* Special categories of cited documents: "A" documents defining the general state of the art which is not considered to be of particular relevance "T" earlier documents published on or after the international filing date "X" documents which may throw doubts on priority claim(s) or which is cited to establish the publication date of another document or other special reason (as specified) "Y" documents referring to an oral disclosure, use, exhibition or other means "Z" documents published prior to the international filing date but later than the priority date defined "a" documents published after the international filing date or priority date and are in conflict with the application but cited to understand the principle or theory underlying the invention "b" documents of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "c" documents of particular relevance, the claimed invention cannot be considered to involve an inventive step if the document is considered with one or more other such documents, each contribution being considered to be a part of the art "d" documents anterior to the same prior art		
Date of the actual completion of the international search 09 NOVEMBER 1999		Date of mailing of the international search report 14 JAN 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20531 Facsimile No. (703) 305-3330		Authorized officer THOMAS LEE Telephone No. (703) 305-9717

Form PCT/ISA/210 (second sheet)(July 1992)*

フロントページの続き

(81)指定国 EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), AP(GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW

Fターム(参考) 5K026 AA21 BB04 CC04 CC08 DD01
EE01 FF07 GG03 HH06 JJ03
5K030 HA08 JA14 JL03 KX20
5K051 AA01 AA03 DD05 EE01 EE07
FF03 KK01
5K069 AA08 AA15 BA01 CA00 CB01
DA02 DA07 DB02 DB04 DB07
DB09 EA30 FA03